# Intro to DNA and RNA

## Introduction

Subsequent lessons will go into intense detail concerning the proteins involved in the processes of replication and transcription. The objective here is to provide a primer and bridge. The information is presented together because both **DNA replication** and **mRNA transcription** rely on **polymerases**—enzymes that copy the DNA by adding a single nucleotide at a time. While there are stark differences between the two processes, the ideas of how a template is read, how a new strand is made, and the directionality with which the polymerases move are quite similar.

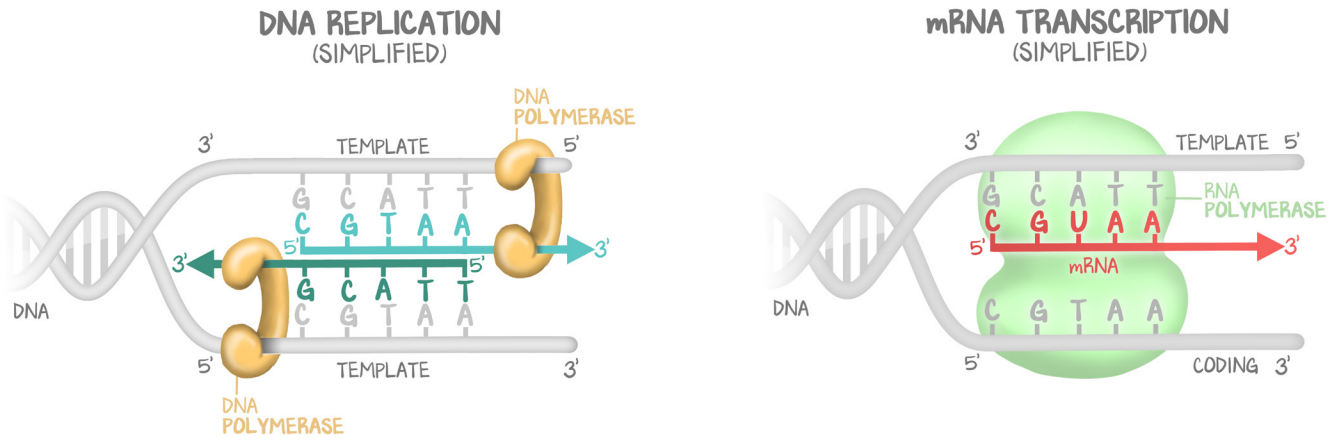


**Figure 4.1: DNA Polymerase and RNA Polymerase**
(Left:) DNA polymerase copies both strands (both template strands) and incorporates DNA. (Right:) RNA polymerase copies only one template strand, the other acting as the coding strand. Both DNA polymerase and RNA polymerase can add to the end of their own 3' strand only.

The **template strand** is the strand of DNA that the polymerase will use as the template to construct the new strand. Whether that new strand is DNA (replication) or RNA (transcription), the polymerase will be doing the same thing.

The **daughter strand** (here referring to either the replicating strand or the transcribing strand) will be constructed **complementary** and **antiparallel** to the template strand. The polymerase will begin somewhere on the template strand and move along it. The polymerase will be moving in the direction of **3' to 5' relative to the template strand**.

This orientation is crucial. The easiest way to keep things straight is to keep the daughter strand **5' to the left of the page, 3' to the right of the page**. Since the daughter strand is made by adding to its own 3' end, the earliest nucleotide will be on the left of the page, with the most recent on the right. That means the template strand, being complementary and antiparallel, is oriented 3' to the left of the page, 5' to the right. Thus, the polymerase moves left-to-right along with the template.

Orientation is more important for transcription than for replication. Replication doubles everything in the genetic code, though the directionality creates the leading and lagging strands. Transcription orients the message in a specific way such that the 5' left to 3' right organizes the mRNA—the coding strand—and the final amino acid sequence in order from left to right, just as we read and write in English.

Nucleotides are added to the end of the growing strand by a **polymerase**. It attaches two nucleotides together, taking the open free 3' end of the already elongated chain, grabbing a free nucleotide, and connecting its 5' end to the 3' of the strand already there. This is done via **phosphodiesterase bonds**. It builds nucleic acid strands in the exact way described in the *Nucleic Acids* lecture. If the daughter strand

is **DNA**, it'll use a DNA polymerase and DNA nucleotides (GC, A**T**). If the daughter strand is RNA, it'll use an RNA polymerase and RNA nucleotides (GC, A**U**).

## Prokaryotic Replication Reveals an Issue with 5' to 3' Directionality: Leading and Lagging Strands

Replication copies a DNA version of EVERYTHING. The process is said to be **semiconservative**. Semiconservative means that half of the DNA sequence (one strand) will end up in the daughter, while half of the DNA sequence (its original paired strand) will end up in the other. If everything goes correctly, DNA is copied identically, so there'll actually be the same DNA sequence in both daughters. Semiconservative means simply that the DNA isn't copied with said copy going into the daughter, but instead the daughter cells each have one half original with one half replicated.

Bacteria have **circular DNA**, with a **single origin site** (Ori C). Proteins identify the origin site, unzip the DNA molecule by breaking hydrogen bonds between base-pairs, and get the cellular machinery to begin to replicate. That replication machine reads **both template strands** at the same time, but **moving in opposite directions** away from each other. Because the original template strands were complementary and antiparallel, and the replicating strand starts 5' and moves to its own 3', both are beginning at their respective 5' and moving in opposing directions to start.

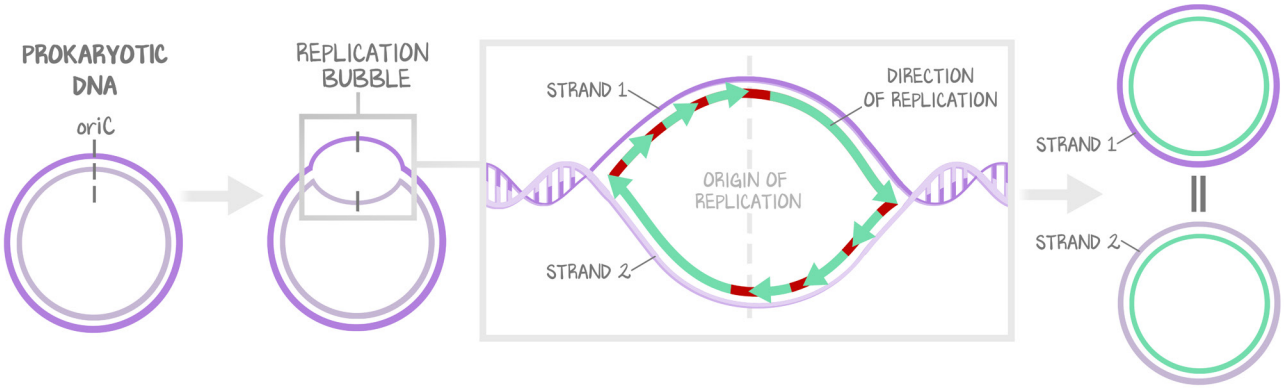


**Figure 4.2: Prokaryotic DNA**
Replication uses a single origin site and introduces the concepts of semiconservative and leading/lagging strands.

One of the polymerases is moving clockwise on strand 1 while the other is moving counterclockwise on strand 2. That means all of the surface area covered by the polymerase on strand 1 isn't interacting at all with the same area over strand 2. And vice versa, the polymerase moving counterclockwise on strand 2 is busy copying strand 2, and ignores strand 1. We could wait for the polymerase to do a complete circle, finishing where it started. That is inefficient, and that's not what happens. In fact, exposing single-stranded DNA like that would only cause it to break, so we have to do something with it.

The **leading strand** is the one being **continuously built**. The **lagging strand** is the counterpart across the gap, filling in piecemeal. The lagging strand is still being built starting at its own 5' moving towards its 3', moving towards the template strand 5'. Every time the DNA reveals what could be a start sequence, a new, short sequence starts and grows towards the original site of the leading strand. But now, because there are intermittent start signals that end with collision with the already duplicated strand, these are added **noncontinuously**. They need to be **stapled together** at the end (the details are in the *Replication Fork* lesson).

Ultimately, the two completely duplicated strands split apart—the original double-stranded pair duplicated. The daughter strands are identical to the original. It's just that each daughter has one strand from the parent, and one new copy.

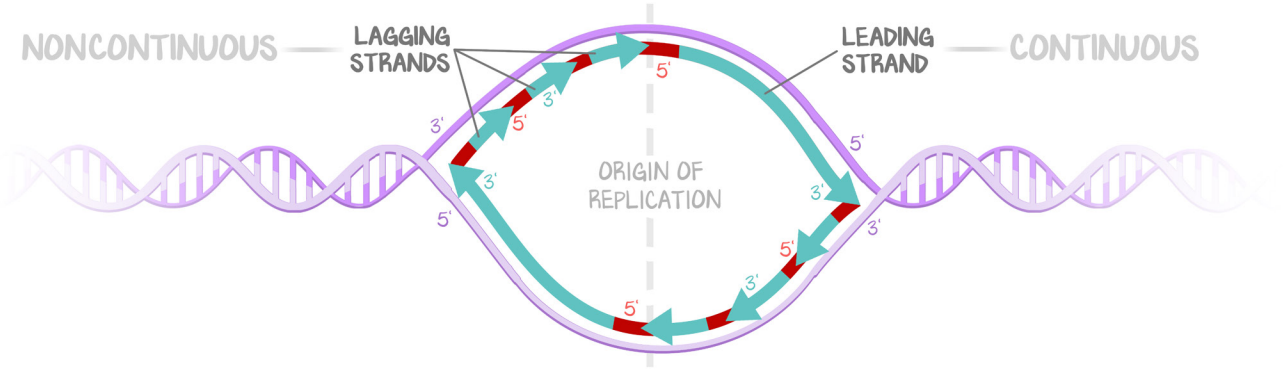


**Figure 4.3: Leading and Lagging Strands**
The origin of replication reveals a continuous uninterrupted leading strand traveling towards the replication fork, and a discontinuous lagging strand building towards the origin on both strands.

## Eukaryotic Replication

Eukaryotic cells do basically the same thing. They split apart, polymerases replicate strands, and the final result is two duplicates of the original, with each daughter having a parent strand and a copied strand. The eukaryotic polymerases start at their own 5' and build towards their own 3', moving towards the 5' of the template strand. The only major difference is that **eukaryotes have multiple sites of origin** and their DNA is not circular, **but instead linear**.

This means the polymerases won't cross each other as in bacteria. Although there are **multiple replication forks**, **each with** its own **leading and lagging strands**, replication will go out to the ends of the chromosome. The two original strands will remain connected at the **centromere**.

In eukaryotes, one linear double-stranded DNA chromosome becomes two linear copies of double-stranded DNA that are connected by a centromere.

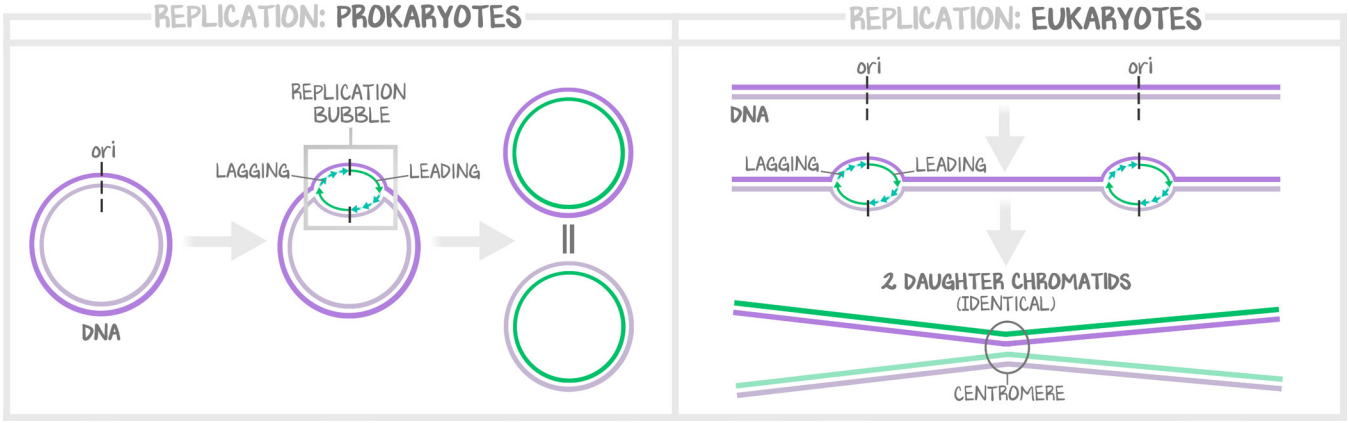


**Figure 4.4: Prokaryotic vs. Eukaryotic Replication**
Prokaryotes copy two strands, have only one ori site, and produce two semiconservative strands. Eukaryotic transcription also produces two semiconservative strands, but is linear and has multiple origin sites.

## Prokaryotic and Eukaryotic Transcription

Transcription, the process of RNA polymerase making mRNA from template DNA, occurs in **one direction** and on **one strand only**. There aren't leading or lagging strands, but instead a **coding** strand.

Said coding strand contains the code, but RNA polymerase never touches it. Master the replication fork (#5 *DNA Replication (The Fork*) and we'll bring it back around to make it easy for mRNA.

**Transcription copies an RNA version of only what's needed right then.**

Prokaryotic replication occurs in cytoplasm, which means it can occur at the same time as translation. Eukaryotic transcription occurs in the nucleus, is packaged, and is then sent to the cytoplasm.

## DNA Polymerase vs. RNA Polymerase

**DNA polymerase** uses a **template DNA** strand to make **another DNA strand**. DNA polymerase is used in **replication**, but not in transcription. The idea of DNA replication carries more weight than transcription. If an error is introduced into the replication process, that error is therefore carried forward in every cell line from that point forward. Those cells vulnerable to this would turn malignant or die. DNA replication, and therefore DNA polymerase, has **more rules and features** than RNA polymerase and transcription. DNA polymerase uses DNA, so the base pairs it knows how to handle are G, C, A, and T.

**RNA polymerase** uses a **template DNA** strand to make **messenger RNA**. RNA polymerase is used in transcription, but not replication. If a bad order goes out from a dysfunctional messenger, the cell can always send another message. And even if a bad message gets through, it affects proteins, not the cell itself, and typically doesn't cause lasting damage to the cell line. Mistakes made by RNA polymerase aren't carried forward through proliferation and replication. RNA polymerase uses RNA, so the base pairs it knows how to handle are G, C, A, and U (uracil for RNA, thymine for DNA).

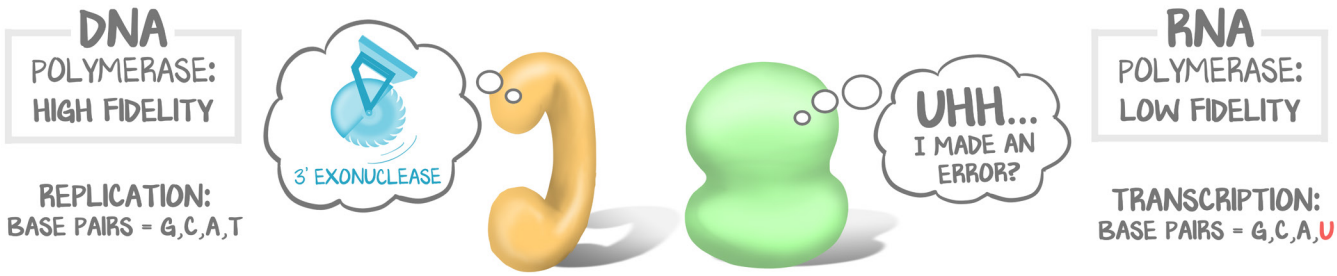


**Figure 4.5: DNA Polymerase vs. RNA Polymerase**
DNA polymerase needs to be high-fidelity to ensure that mutations are not introduced into the cell line, so has an exonuclease repair mechanism and proofreading, whereas RNA polymerase does not. They differ in their handling of nucleotides only in that T is DNA while U is RNA.

**DNA polymerase** and **RNA polymerase** both move from their own 5' to their own 3', moving towards the 5' of their template strand.

As mentioned, since DNA polymerase has a more serious job to do (on longevity of the cell line), DNA polymerase has a few more rules and features. DNA polymerase **requires an RNA primer** to ensure that it's starting at the right spot. This is true of the leading and lagging strands. DNA polymerase is **high-fidelity**. Energy is spent on ensuring that the copies are good. If there's an error, even during the process of replication, the DNA polymerase can fix it. DNA polymerase has a **3'→5' exonuclease** that allows the DNA polymerase to back up, remove a bad pair, and continue forward via a process called **proofreading**.

RNA polymerase doesn't require a primer, elongates only (the mRNA and the DNA strand don't need a hydrogen bond), and has no exonuclease so can't proofread. Thus, RNA polymerase is said to have **low fidelity** (the end product has errors, and it's okay . . . usually).